

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2001, 19:40:41 ; Search time 181.41 Seconds

(without alignments)
18724.010 Million cell updates/sec

File: US-08-153-397A-1

Perfect score: 3962

Sequence: 1 CGGCGCTGACGACGCGTGA.....AAAAAAAAAACCGGAATTC 3962

Scoring table: IDENTITY_NUC 11
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: /SID2/gcgdata/geneseq/geneseq/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3962	100.0	3962	AA93785	Human mammary carc
2	3960.4	100.0	3962	AA92520	Human mammary carc
3	3960.4	100.0	3962	AA92522	Human mammary carc
4	3617.6	91.3	3754	AA084782	Human mammary carc
5	2783.6	70.3	2861	AA09051	Protein-tyrosine-k
6	2783.6	70.3	2861	AA09051	Human pancreatic c
7	642	16.2	3096	AA48292	Human colon cancer
8	642	16.2	3157	AA092521	Discolidin domain r
9	642	16.2	3157	AA092523	Human colonic aden
10	642	16.2	3157	AA092523	Human colonic aden
11	639.8	16.1	3120	AAV5895	Receptor protein t

12	639.8	16.1	3120	AAV5895	Receptor protein t
13	544.2	13.7	2725	AAH99505	Human protein enco
14	453.4	11.4	2128	AA064158	Partial coding seq
15	324.4	8.2	408	AA000624	Human secreted pro
16	236	6.0	272	AA022477	Human gene signatu
17	228.4	5.8	463	AA11731	Probe #1664 for ge
18	228.4	5.8	463	AA13038	Probe #1724 used t
19	228.4	5.8	463	AA101659	Probe #1650 used t
20	182.2	4.6	2820	AAV51456	gd.trkb fusion use
21	180.6	4.6	2301	AAV20445	Human c-trk oncoge
22	180.6	4.6	2301	AAA09303	Human trk oncogene
23	180.6	4.6	3060	AAV51457	gd.trkb fusion use
24	180.6	4.6	3194	AAV51457	Human trkb recepto
25	180.6	4.6	3194	AAV51457	Human trkb recepto
26	180.6	4.6	3707	AA099277	Human neurotrophic
27	180.2	4.5	2526	AA069029	Porcine trk gene.
28	178.6	4.5	2526	AA028668	Encodes adult porc
29	178.6	4.5	2526	AA034581	trkC clone in pFLI
30	178.6	4.5	2940	AAV51458	gd.trkb fusion use
31	177	4.5	2750	AAV51458	DNA coding for neu
32	158	4.0	4092	AAV70230	Probe #10877 for g
33	151	3.8	175	AA10944	Human receptor tyr
34	151	3.8	175	AA16188	Probe #14874 used
35	151	3.8	175	AA106655	Probe #6646 used t
36	140	3.5	2376	AA034582	Partial sequence o
37	139	3.5	2208	AAV16347	Mouse muscle-local
38	139	3.5	2604	AAV16349	Mouse muscle-local
39	138.4	3.5	2376	AA028669	Encodes mouse trkC
40	138.4	3.5	2376	AA069030	Murine trkC gene.
41	138	3.5	3900	AA068343	Degenerate rat ins
42	137.4	3.5	3257	AAV14352	Nsk2 receptor gene
43	137.4	3.5	3257	AAV18546	Mouse receptor tyr
44	135.2	3.4	1212	AAV59295	Human insulin rece
45	135.2	3.4	4723	AAH50568	Insulin receptor c

ALIGNMENTS

RESULT 1

AA93785 standard; CDNA: 3962 BP.

AA93785	16-FEB-1998 (first entry)	Human mammary carcinoma kinase 10 (MCK-10) cDNA sequence.
AA93785		Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase; proliferative disease; cancer; insulin receptor family; tyrosine kinase neurotrophin receptor; MCK-10 activity; neurological disorder; aberrant expression; ds.
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	321..3077
FT		/*tag- a
XX		
FN	US5677144-A.	
XX		
PD	14-OCT-1997.	
XX		
PF	08-NOV-1994; 94US-0336343.	
XX		
PR	16-NOV-1993; 93US-0153397.	
XX		
PA	(ALVE/) ALVES F H E.	
XX	(ULR/) ULLRICH A.	
XX		
PI	Alves FHE, Ullrich A;	
XX		

QY 1741 ACCGCCAGTCTAGAGAGCCACCCGCTACAGAGGCCCGGCTGTGGAAATCCGC 1800
|||||
Db 1741 aacgcgcaggtcctctagagagccaccccgctacagagagcccggtctctggaatccgc 1800
QY 1801 CCCACTCCGCTCCCTGTGTGCCCAATGGCTCTGCTGTGCTCTCAATCCAGCCTTAC 1860
|||||
Db 1801 cccaatccgcctctgtctcccaatgctctcgctgtctgtctcccaatccagctaac 1860
QY 1861 GCTCTCTTGTGGCCACTTACGCCCTCTCCCTCGAGGCCCGGCCCCCACACCCGCT 1920
|||||
Db 1861 gctctctctgtgccaactacgcccgtccctctgagcccggtccccccacacccgct 1920
QY 1921 GGGCGAAACCCACCAACCCAGGCTACAGTGGGAGTATATGAGAGCTTGAAAGCCAG 1980
|||||
Db 1921 gggccaaaacccaacacccacagagctacagtgggagctataatgagctggaagccag 1980
QY 1981 GCGGCCGCTTCTGCCCCCACCCTCCCGAGAACAGCGTCCCATATGCGGAGGCTGACA 2040
|||||
Db 1981 ggcgcgcgctctgtcccccaactcccaagacaggtcccatatgtcgaagctgaca 2040
QY 2041 TTGTATCCCTGTCAGGGGCTCACCGGGGGCAACACTATGCTGTGCTGCTACCTGCCAG 2100
|||||
Db 2041 ttgtacctgtcaggggtcacccggggcaacactatgtgtgtgctgacctgccccag 2100
QY 2101 GGGCAGTCCGGGATGGGCCCCCAGAGTGGATTTTCCCTCGATCTCGACTCCGCTTCAAG 2160
|||||
Db 2101 gggcagtcggggagtgggccccccagagtgatctccctcgatctcgactccgctcaagg 2160
QY 2161 AGAGCTTGGCGSAGGGCCAGTTTGGGAGTGCACCTGTGAGGCTGCACAGCCCTTCAG 2220
|||||
Db 2161 agaaagcttgcgaggggccaagtcttggaaggtgcacctgtgtgaagtgcaagccctcag 2220
QY 2221 ATCTGTCACTTGTATTTCCCTTAATGTGCTTAAGGAGCACCTTGTGTGAGTGTG 2280
|||||
Db 2221 atctgtcagctctgtatctcccttaattgtgcttaaggagacacctgtgtgtagtgtg 2280
QY 2281 TCAAGATCTTACGGCCAGATGTCACCAAGAAATGCCAGCTTCTCTTGTTCAGGAATG 2340
|||||
Db 2281 tcaagatcttacggccagatgccaacgaagatgcagctctctctgtctccsagaatg 2340
QY 2341 ATTTCGAAAGAGGTGATGATGTCGAGGCTCAAGAGCCCAACATCATTTGGGCTGC 2400
|||||
Db 2341 atttcctgaagaggggtgaataatcatgtctcgaaggtcccaagacccaatctctgctgc 2400
QY 2401 TGGCGCTGTGTGTCAGAGACGCCCTCTGTGATGATTACTGATACATGAGAGAGCGC 2460
|||||
Db 2401 tggcgctgtgtgtgacagagacgccctctgtgcatgattactgactaataatgagaagcgcg 2460
QY 2461 ACCTCAACAGTTCTCTAGTGCACACAGCTGAGAGCAAGAGCAGCGAGGGGCCCTG 2520
|||||
Db 2461 acctcaacagttctctcagtcagccaacagctggaagcagcgagggggccctg 2520
QY 2521 GGGACGGGCGAGGTGGCGACAGGGGCCACATACGCTACCCCAATGCTCTGTGATGTGGCAG 2580
|||||
Db 2521 gggacgggcgaggtgtgcagaggggcccaatcagctacccaatgtctgcatgtggaag 2580
QY 2581 CCCAGATCCGCTCCGSCATGCGTATCTGGCCACACTCAATTGTATCATCGGAGCTG 2640
|||||
Db 2581 cccaagtctgcctcgcgcatgtctatctgtgccacactcaactctgttcaatccggaaactg 2640
QY 2641 CCACGGGGAAGTCTGCTAGTTGGGAAATTTCAACATCAAAATCGACAGCTTTGGCATGA 2700
|||||
Db 2641 ccacggggaagtctcctagttgggaaaatttcaacatcaaaatcgcagactttgcatga 2700
QY 2701 GCGGGAACCTTATGCTGGGAGTATTTACGCTGTGACGGGCGGCGAGTGTGCTGCCATCC 2760
|||||
Db 2701 gccggaaacctctatgtctggtggaactataacggtgtgcaagggcggaaggtctgccccctcc 2760
QY 2761 GCTTGATGGCTGTGGAGTGCATCTCATGGGGAAGTTTCAAGACTGCGAGTGTGTGGG 2820
|||||
Db 2761 gcttgatggcctgtggagtgatctcatctcatgtggaagttaacgactgtgagtgaaagtgtggg 2820

QY 2821 CCTTGTGTGACCTCTGTGGAGCTCTGATGCTGTGTAGGGGCCACCCCTTTGGGACG 2880
|||||
Db 2821 ccttgtgtgtgacctctgtgtggagctctgtatgtctgtatggtccagcccttgggcagc 2880
QY 2881 TCACGAGAGAGAGTCAATCGAAGACGGGGAGTTCTTCCGGGAGCCAGGCGGCGAG 2940
|||||
Db 2881 tcacgagagagaggtcaatcgagaacgctgggagttcttccgggacagggccggcagag 2940
QY 2941 TGTACTGTCCGGCGCCCTGCTGCCCCCAGAGGCTATATGAGCTGATGCTGGTGTCT 3000
|||||
Db 2941 tgtactgtccggcgccgctctgtccgcgcaaggctctataatgactgtactgtctgggtct 3000
QY 3001 GGAAGCGGAGTGTGAGAGCGACCACTTTTCCAGCTGATGCTGTCTCTGAGAGAG 3060
|||||
Db 3001 ggaagcggagtgctgagcagacaccccttctccagtgatcgtgtctccgggaagag 3060
QY 3061 ATGCACTCAACAGGTTGTAATCACATATCCAGCTGCCCTCTCCAGGAGTATCCAG 3120
|||||
Db 3061 atgcaactcaacaggtgtgaatcacatccagctgcccctccctcagggagtacag 3120
QY 3121 GGAAGCCAGTGAACCTAAACAGAGACAAATGAGGACGCTGCGCTTCCCTCCCGGA 3180
|||||
Db 3121 ggaagccagtgtaacctaaacagagacaaatgagcaatctgacctctccctcccgga 3180
QY 3181 CAGCCATCACCTTAATAGAGAGAGTGAAGTGAAGTGGGCTGGGCCACCAAGAGAG 3240
|||||
Db 3181 cagcccatcaaccttaatagagagcagtgaactgcaagtggtgtgtggcccaacagggag 3240
QY 3241 CTGATGCCCTTCTCTCCCTTCTCTGACACACTCTCATGTGCCCTTCTGTCTCTTCC 3300
|||||
Db 3241 ctgatgcccttctctcccttctctgacacactcatgctccctctcctgtctctctcc 3300
QY 3301 TAGAAGCCCTGTGCGCCACCCAGCTGCTGCTGTGAGTGGATTCCTTCACACCTCTCT 3360
|||||
Db 3301 tagaagccctgtgctgcccccaacagctgtccctgtgagtggtctctccacacctctct 3360
QY 3361 AGCCATCCCTTGGGAGAGGTGGGGAATATAGATAGACACTGACATGAGCCCATTTG 3420
|||||
Db 3361 agccatcccttggggaaggggtgggaataataggaatagacactggaatgagcccatgtg 3420
QY 3421 GAGCACTGGGCCCCACTGACACACACTGATCTCTGAGAGGTGGCTGGCCCCAGCTTC 3480
|||||
Db 3421 gagaacctgggcccccaatgcaacactgattctctggagaggtgtgctgcgccccagcttc 3480
QY 3481 TCTCTCCCTGTACACATATGAGACCCCACTGGCTGAGAACTCTGGGGGTAGAGAGCAAGA 3540
|||||
Db 3481 tctctccctgtcacacatctggaaccccactgtgtagaaatctgggggtgtaggaagacaaga 3540
QY 3541 AGGAGAGGAAATGTTCCTTGTGCTGCTCTCTGTAATTTGCTCAAGCTTGGGCTTCTC 3600
|||||
Db 3541 aggagaggaataatgttctctgtgtcctgtcctctgtactgttcccaagcttgggtcttc 3600
QY 3601 CTCCTCATCACTGAAACACTGACCTGGGAGTACCCGCCCAAGCCCTCACTCAACC 3660
|||||
Db 3601 ctctcatcaatcctgaaacactgtgacctgggtgtagccccgcccaagccctcagctaaccc 3660
QY 3661 CCACCTCCCACTTGAGCTGTAGCTGTAGTAACCTTCTTAACCTTATAGTGTTCGTGGAG 3720
|||||
Db 3661 ccacttcccaactgtcagctgtgtagaaactctctcgaacccatacgttctctgtgag 3720
QY 3721 TAAATATTGGATTGGGGGAAAGAGGAGAGCAAGGCCCAATGAGCTTGGGGTGGACATC 3780
|||||
Db 3721 taatatttggaattggggggaagaggaaggaagcccaatagccttgggggttggaaatc 3780
QY 3781 TCTAGTATAGTGCACATTTGATTTTCTATATCACTTGGGGTGTATATTTTGGGG 3840
|||||
Db 3781 tctagtgtagctgcacatgtatcttctataatcaacttgggtgttctacatcttttgggg 3840
QY 3841 GGAAGAGACAGATTTTATACATTAATATGACCTAGCTGTGAGGCAATTTATCCCT 3900
|||||
Db 3841 ggaagagacagatcttctacataatataatgacctagcttggaggaatcttaatccct 3900
QY 3901 GCACTAGCAGGTAAATAAAGTTTCCACAAAAAATAAATAAATAAATAAATAAATAA 3960


```
Db 241 tcaactgacgatagggttgacttgaaagaaatgccaagatgctgcaccccaacccctta 300
QY 301 GGGCCGAGGAGTCAAGGAGCTATGGAGACGAGCCCTGTCACTTATCTTACTGCTGCTCT 360
Db 301 gggccgagggatcagagaaactatggaacagagccctgtcaacttactgctgcctct 360
QY 361 TGTGTGCAAGTGGAGATGCTACATGAGAGACATTTGATCTGTCCAAAGTGGCCGTATG 420
Db 361 tgggtgcaagtgagatgctgactgaaagaaatggaacatttgactctgcacagtcgctatg 420
QY 421 CCCGTGGGATCAGAGACCGGACCATCCAGACATGACATCTCTCTTCCAGCTCTGTGT 480
Db 421 cccgtgggatacagagaccggaacatcccaagacagtgacatctctctccagctcctgt 480
QY 481 CAGATTCCACGCGCCCGCCGACACbbaagttggagagacagatgacgggagtgggcctgt 540
Db 481 cagattccacgcgcccgcccgacacaggttgagagacagtgacgaggatgggagcctgt 540
QY 541 GCGCCGAGGGTCCGATTTCCCAAGGAGGAGGATCTTGAGAGTGGATGATACAAACAC 600
Db 541 gcccgcagggctcggttctccaaagagagagatctgcaggtggacttaacacagac 600
QY 601 TCCACGTGTGGCTCTGTGGTGGGACCCAGGAGCGCATGCGGGGCGCTGGGCAAGAGT 660
Db 601 tccacgtgtgctctgtgtgggacccagagacgcatgcccggggccttggcaagagat 660
QY 661 TCTCCCGGAGTACCGGCTGCGTTACTCCGGGATGTCGGCGCTGATGGGCTGGAAG 720
Db 661 tctcccggaactacgctgctgtactccggagtgtgcgcgctggatgggtggagag 720
QY 721 ACCGCTGGGTCAGAGAGTGTATCTCAGGCAATGAGGACCCGAGGAGATGGTGCTGAAG 780
Db 721 accgctgggtcagagagtgatctccaaagcaatgagaccgagagatgggtgctgaag 780
QY 781 ACCTTGGGCCCCCATAGTTGCCCCGCACTGGTCCCTTCTACCCCGGCGCTGACCGGTC 840
Db 781 accttgggcccccatagtgtggccgactgtctcttaccgcccggtctaacgggtgca 840
QY 841 TGAATGCTGTCTGCGGGGATGAGCTCTATGAGCTGCTCTGAGGAGATGAGTCTGTCT 900
Db 841 tgaatgctgtctgcggggatgagctctatgagctgctcttggagagatgactcctgt 900
QY 901 ACACCGGCGCGTGGGGGACAGAAATGATTTATGTAGAGCGCGTGTACTCAACACTCA 960
Db 901 acacggcgccgtgggggacaaatgatatctatctgagccgtgtlaccatcaacga 960
QY 961 CCTATGACGAGACATACCGTGGGCGAGCTGCAGTATGGGGGCTGAGCGTGGCAGATG 1020
Db 961 cctatgacgaaacataccgtgtgggactgcagtaatgggggtcgtggcagactggcagatg 1020
QY 1021 GTGTGGTGGGCTGATGACTTTAGAAAGTCAAGAGCTGCGGGCTGAGCGAGCTATG 1080
Db 1021 gtgtgggtgggctgatatgactttagaaagtcaagagctgcgggtcttgcagagctatg 1080
QY 1081 ACTATGTGGATGAGACAAACACAGCTTCTCCAGTGTGTGAGATGAGATGATTTGAGT 1140
Db 1081 actatgtggatgagacaaacacagcttctccagtgtgctatgtgagatgagattgagt 1140
QY 1141 TTGACGGGCTGAGGGGCTTCCAGGCTATGAGGTCACGTAAACAATGACACGCGTGG 1200
Db 1141 ttgacgggctgaggggcttccaggtcatgcaagttccacgtlaaacaatgacacgctgg 1200
QY 1201 GAGCCGCTGCTGCTGGCGGGGTGAATGTGCTTCGCGCGTGGCGCTGCCATGGCTGGG 1260
Db 1201 gagccgctgctgctggcggggtgaaatgtgcttcgscggtggccctgcattggccggg 1260
QY 1261 AGGGGAGGCGCATGCGCCACACACTAGAGGGGCAACTGGGGGACCCAGAGCCCGGCTG 1320
Db 1261 aggggagggcgcatgcgccacacactaggggcaacctgggggagcccgagcccggtcg 1320
QY 1321 TCTCAATGCCCTTGGCGGCGGTGTGCTGTGCTTGTGAGTGCCTGCTTTCGGG 1380
Db 1321 tctcaatggccttggcgcggtgtgctgtgcttctgcatgagtcgcgcttcttgg 1380
QY 1381 GGGCCTGTGTACTCTTCAAGCAAAATCTCTTATCTGTGATGTGGTGAACAAATCTCTTC 1440
Db 1381 gggcctgtgtactcttcaagcaaatctcttcaatctctctgattgtggaacaattctctcc 1440
QY 1441 CGGCACTGGGAGGCACTTCCCGCCAGCCCGCTGGTGGCGGCTGGCCCACTCCACCA 1500
Db 1441 cggcactgggagggcaacttcccgccagcccgctgggtggcgccctggcccaactccaca 1500
QY 1501 ACTTCAGCAGCTTGGAGCTGAGAGCCAGAGGCGAGAGCCGCGGCAAGGGCGAGGGA 1560
Db 1501 acttcagcagcttggagctgagagccagagagccagagccgttgcgaagggcgaggga 1560
QY 1561 GCCCGACCGCATCTCATCGGCTGCTGCTGATGAGCATCACTATCCCTCATCACA 1620
Db 1561 gcccgacccgatctctcatcggtctgctgtgtggatcaatctctctcctgcgtcatca 1620
QY 1621 TTGCCCTCATGCTGTGGCGGCTGACATGGCGCAGGCTCCTCAGCAAGGCTGAACGAGG 1680
Db 1621 ttgccctcatgctgtggcggtgacatggcgacggctcctcagcaaggctgaacgaggg 1680
QY 1681 TGTGGAAGAGAGGTGAGCGGTTCACTCTGCTGCTGGGGAACATATCCCTCATCACA 1740
Db 1681 tgtggaaagagagtgagcggttcaactctctgtcccttgggaacatctcatcaaca 1740
QY 1741 ACCGCGCAGGTCCTTAGAGAGCCACCCTACAGAGAGCCCGGCGCTGTGGAAATCCGC 1800
Db 1741 accgcgcggtccttagagagccaccctacagagagcccggtcgtggaaatccgc 1800
QY 1801 CCCACTCGGCTCCCTGTGTGCCAAATGGCTCTGCGTGTGCTTCATTCACGCTTACC 1860
Db 1801 cccactcggctccctgtgtgccaaatggctctgctgtgcttcacatccagctacc 1860
QY 1861 GCCCTCTTGCGCACTTACAGCGCGCTCCCGCCGAGGCGCCGGGCCCCCACAACCGGCT 1920
Db 1861 gccctcttgcgcaacttacagcgctcccgccagggcccgagcccccacaacggcct 1920
QY 1921 GGGCCAAACCCACAAACCCAGGCTACAGTGGGAGACTATGTAGAGCTGTGAGAGCCAG 1980
Db 1921 gggccaaacccacaaacccaggtacagtgggagataatgagctcgagagagcag 1980
QY 1981 GGGCCCGCTCTGCGCCCACTCCCAAGACAGCGTCCGCCATTTATGCCAGGCTGACA 2040
Db 1981 gggcccgctctgcgcccaactcccaagaaaggtccccaattatgcagaggtgaca 2040
QY 2041 TTGTATACCTGTCAGGGGCTCACCGGAGGCAACCTATGCTGTGCTGACTGCCCCAG 2100
Db 2041 ttgtatacctgtcaggggtcacccggggcaacaactatgtgtgctgtcacgtgccccag 2100
QY 2101 GGGCAGTGGGAGTGGGCGCCCAAGAGTGGATTTCCCTCGATCTGAGTCCGCTTCAAG 2160
Db 2101 gggcagtgggagtgggcgcccaagagtggatcttcccagatctgcagctccgctcaag 2160
QY 2161 AGAAGCTTGGGAGGCGCAGTTTGGGAGTGCACCTGTGTAGAGTGCAGACCCCTCAAG 2220
Db 2161 agaagcttgggagggcgcaagttggggagtgcacctgtgtgaggtgcgaagccctcaac 2220
QY 2221 ATCTGGTATGCTTGTATTTCCCTTAATGTGCGTAAAGGACbbaoccttggctgtagctg 2280
Db 2221 atctggatgcttgtatttcccttaattgtgctlaaggaacaccccttgtgtagctg 2280
QY 2281 TCAAGATCTTACGGGCGAGATGCCACAAGATGGCAGATTTCTCTGTGTCGAGGATG 2340
Db 2281 tcaagatcttaccggcgagatgccacaagatggcagatcttctctgttcccgagatg 2340
QY 2341 ATTCTGAAAGAGGTGAAGATCATGTGAGGCTCAAGGACCCCAAGATCATTTGGGCTGC 2400
Db 2341 attctgaaagaggtgaagatcatgtgaggtcctcaagaccaccaatacttgcgtgc 2400
QY 2401 TGGGCGTGTGTGTGTCAGAGAGACCCCTCTGTGATGATTTACTGACTACATGGAGAGCGG 2460
Db 2401 tgggctgtgtgtgtcagagagacccctctgtgactgattactgactacatgagaaaggcg 2460
```

OY	2461	ATCTAACACAGCTTCCTCAAGTGGCCCAACCAAGCTGGAGAGACAAAGGACGACGGAGGGGCCCTTG	2520
Db	2461	accctcaacacagcttctctcaatgctcccaacacagcttggaaagcaagcgacgacggaggggccccctg	2520
OY	2521	GGGACGGGAGGGCTGCGCAGAGGGGCCCACTACAGCTCCCAATGCTCTCTCATGTGGCAG	2580
Db	2521	gggaacgggacagctctgctcgcaaggggcccacacctcacgctaccccaatgctctgctcaatgctgcaag	2580
OY	2581	CCCAAGATGGCTTCGGGATGCGGCTATCTGGCCACACTCAACTTTGTATCATCGGGAGCTTGG	2640
Db	2581	cccaagatgctctccgcgcatgctgctatctctggccacactcaaccttctgctacatcggaactctg	2640
OY	2641	CGACGGGGAACTCTCCTTAAGTTGGGGAAAATTTCACATCAAAATAATGCAAGACTTTTGGCAATGA	2700
Db	2641	ccacggggaaactctccatgcttggggaatttcacatccacataatcgcaacttgcagacttgcata	2700
OY	2701	GGCGGAACCTATATGCTGGGGGCAATTAATACGTGTGGACGGGCCGGGAGCTCTGCCATTC	2760
Db	2701	ggcgggaacctctctatgctctggggactatctaacgcgtctgcagggcgcgggcgacgtctgcacatcc	2760
OY	2761	GCTGTATGGCTTGGGAGAGTGCATCTCATGGGGAAATTCACAGATGGCGAGTGGAGTGGGG	2820
Db	2761	gctgtatggcttgggagagtgcacatctctcatctcattgggaagtcaacgactgagatgtaagctgtggg	2820
OY	2821	CCCTTGTGTGACCCCTGTGGGAGGTGCTGATGCTCTGTAGGGGCCACGCCCTTTGGGCAAGC	2880
Db	2821	cccttgtgtgacccctgtgggaggtgctgatagtctgtatgagggccacgccccttctgggacgc	2880
OY	2881	TCACCCACAGACAGGTCATCTGAAAGAGCGGGGAGTCTTTCGGGGACACAGGGCCGGCAGG	2940
Db	2881	tcacccacagacaggtcatctgaaagagcggggagttcttcgggacacagggcgagggcgaggg	2940
OY	2941	TGTACTGTCTCCGGCCGCTGCTGCTGCCGACAGGGCCCTTATATGAGCTGATCTTCGGTGCT	3000
Db	2941	tgtactgtctccggcccgctgctgctgcgcgacgggacctatataatgctcgtatgctctggtctgct	3000
OY	3001	GGAGCGGGGAGTCTAGAGCAGACCAACCCCTTTCCACGCTGCATCGGTTCTTGGCAGAG	3060
Db	3001	ggagcggggagttctagagcagacacaccccttccacagctgcatacggtcttcctgacagagg	3060
OY	3061	ATGCATCTAACAGGGGTGAATCACACATCAAGCTCCGCTCCCTCATAGGGAGTGAATCCAG	3120
Db	3061	atgcactctaaacaggtgtgtatcatcaacacacagctctgcccctccctcaaggagtgatccag	3120
OY	3121	GGGAAAGCCAGTACACTAAACAAAGAGACACAAATGGCACTCTGCGCTTCCCTCCGCA	3180
Db	3121	gggaaagccagtacaactaaacaaagagacaaatgagcacctctgccttcccctcccgaa	3180
OY	3181	CAGCCCATCTACTCTTAATAGAGGCGATGAGACTGCAGAGTGGGCTGGGCCACCCAGAGAG	3240
Db	3181	cagcccatcactcttaataagagcgatgtgagatctgacgttgggcttgggcccacccagggag	3240
OY	3241	CTGATGAGCCCTTCTCCCTTCTCTGGACACATCATATGCTCCCTTCCTCTCTCTCTCTCC	3300
Db	3241	ctgatgagcccttctcccttctccctctctggaacaaactctcaatgccccctctgcttcttccctcc	3300
OY	3301	TAGAAAGCCCTGTGCGCCACCCAGCTGTGCTGTGGATGGAGTATCTCTCATCCACTCTCT	3360
Db	3301	tagaaagccctgtgcgcccacccagctgtgctgtgtatgggactctcttccacacccctctct	3360
OY	3361	AGCCATCTCCTTGGGGAAGGTTGGGGAAGAAATATAGATATGACATGACATGGACATGGCCCATTTG	3420
Db	3361	agccatctcttggggaagggttggggaagaataatagatatagacatctggaatgagcccatgtg	3420
OY	3421	GAGCACTGTGGGGCCCAAGAGCAACATGATTCTCTGGAGAGGTGGCTGCGCCCAAGCTTC	3480
Db	3421	gagcactgtggggcccaagagcaaacatgattctctggagaggttgcctgcgccccacagcttc	3480
OY	3481	TCTCTCCTGTACACATGTGACCCCACTGGCTGTGAGATCTGGGGGTATGAGGAGGACAAAGA	3540
Db	3481	tctctcctgtcacacatgtgaccccaactgctctgagaaatctcgggggtgtgaaggaagagcaaga	3540

QY	3541	AGGAGAGGAAATGTTTCCTTGACGCGCMCTGTACTTGCTCCACACTGGCGTTCTTC	3600
Db	3541	aggagaggaaaatglttctcttgcgcctccgtacttgcctcagcttggctcttc	3600
QY	3601	CTCCTCCATCACCCTGAAACACTGGACCTGGGGGTAGCCCGCCACACCTTCAGTCACC	3660
Db	3601	ctctccatcacctgaaacactggaacttgggtggagcccgccccagcctctgctacc	3660
QY	3661	CCACTTCCCACTTGGCAGTCTTGTAGCTAGAACTTCTCTAAGCCTATACGTTTCTGTGGAG	3720
Db	3661	ccacttcccacttgcagctcttctgactagaaactctctaaagctatacgttctctgtgag	3720
QY	3721	TAAATATTGGATTGGGGGAGAAAGAGGACCAACGCCATAGCCTTGGGGTTGGACATC	3780
Db	3721	taataatggtgatctgggggaaagaggagcaacgcccctaaaccttgggttctggaatc	3780
QY	3781	TCTAGTGTAGCTGCCACATTTGATTTTCTATATAATCACTTGGGGTTGTACATTTTGGG	3840
Db	3781	tctagtgtagctgcacatgtaatttctctataacacttgggttctgtacatcttggg	3840
QY	3841	GGAGACACACAGATTTTTCACATAATATAGACATGCTTGAGGCAATTTATCCCT	3900
Db	3841	ggagagacacagatcttctcaacaataatagacacagcttggagcaatcttaatccct	3900
QY	3901	GCACATGAGCAGGTAATATATAAGGTTGAGTTTCCACAAAAAATAAAAAACCGGAAT	3960
Db	3901	gcacatgagcaggtaaataataaggttcttccaaaaaataaaaaaacccggaat	3960
QY	3961	TC 3962	
Db	3961	tc 3962	

XX	AA084782	standard; DNA; 3754 BP.
XX	AA084782;	
XX	17-AUG-1995	(first entry)
XX	Protein-tyrosine-kinase	PTK22.
XX	Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;	
XX	breast tumor; mamma carcinoma; diagnosis; prognosis; therapy; ds.	
XX	Homo sapiens.	
OS	Key	Location/Qualifiers
XX	CD5	142.2886
XX	FT	/tag-a
XX	W09502187-A.	
XX	19-JAN-1995.	
XX	08-JUL-1994;	94MO-GB01480.
XX	09-JUL-1993;	93GB-0014271.
XX	(CANC-) CANCER RES INST.	
XX	(WELL) WELLCOME FOUND LTD.	
XX	Barker RT, Crompton MR, Gusterson BA, Martindale JE;	
XX	Pi Mitchell PJ, Page MJ, Spence P;	
XX	MP1; 1995-066991/09.	
XX	P-PSDB; AAR71100.	
PT	Method for screening substances, using protein tyrosine kinase -	
PT	for potential utility as therapeutic agents for cancer	
DS	Claim 1; Page 26-30; 51pp; English.	

Claim 1; Page 26-30; 51pp; English.

XX cDNA derived from tumor metastatic tissue was amplified using
CC primers (given in AA084783-84) based on sequences (AA71101, AA71103)
CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was
CC identified in an isolated subclone. The 3' sequence of PTK22 was
CC obtained by reverse transcription (using the primer of AA084786) and
CC PCR amplification (primers AA084787-88) of RNA of human breast
CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22
CC is given in AA084782.

XX Sequence 3754 BP: 713 A; 1145 C; 1121 G; 775 T; 0 other;

1 Query Match 91.3%; Score 3617.6; DB 16; Length 3754;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3728; Conservative 0; Mismatches 19; Indels 29; Gaps 7;

QY 181 CCCGGTCGAGCGCTGGGTCTGCCGGAGAGCGATGAGAGTGTCTGAAGTGGCTAT 240
DB 2 ccggctcgagccgctgctcgcgggaagagcagatgaggggtgctcgaaggtgctcat 61
QY 241 TCACTGAGCGATGGGGTGGACTTGAAGAAATGCCAAGATGCTGCCCAACCCCTTA 300
DB 62 tcactgagcgatgggggtgctgacttgaagaaatgcgaagagtgtgccccaccctta 121
QY 301 GGCCTGAGGATCAGAGCATATGAGACAGAGCCGTCTCTTTACTGCTGCTCT 360
DB 122 ggcctgagagatcagagcatatggagacagagccgtcactccttactgtctgctcct 181
QY 361 TGGTGGCAATGAGATGATGATAGTAAGGACATTTGATCTCTCCAAAGCCGCTATG 420
DB 182 tgggtgcaatgagatgatatgataaggaacatttgaatcctccgaagtgcgctatg 241
QY 421 CCTGGGCAATGAGACAGCGAATCCCAAGATGATCTCTCTCTCAAGCTCTGCT 480
DB 242 cctgggcaatgagacagcgatcccaagatgacatctctctcgaagctcctgct 301
QY 481 CAGATTCATGTCGCGCCGCGACAGCGTGGAGACCATGACGGGAGATGGGCTCTGT 540
DB 302 cagattcactcgcgcgcgcacacagcagcttgagagcagtgacggagggcctgct 361
QY 541 GCCCGCAGGGTGGTGTTCCTCCAGAGAGAGAGTACTTGAAGTGGATCTCAACGAC 600
DB 362 gcccgacagggctggtgttcccaagagagagatccttgcaggtggaactcaagac 421
QY 601 TCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 422 tccactgattgattgattgattgattgattgattgattgattgattgattgattgatt 481
QY 661 TCTCCGAGGATACGGGCTGCTACTCCCGGATGCTGCCCTGGATGGGCTGAAG 720
DB 482 tctccgagagatcagggcgtactcctccgcatgtgcgcgtgattgattgattgatt 541
QY 721 ACCGCTGGGAGAGAGTATCTAGCAATGAGACCTGAGGAGAGTGGTCTCAAG 780
DB 542 accgctgggagagagatctagcaatgagacctgagagagagagagagagagagag 601
QY 781 ACCCTGGGCCCCCATGCTGCCGATGCTTCAACCCCGGAGGCTGACCGGGCTCA 840
DB 602 accctgggcccccatgtgtccgactgttgccttgcacccccgggttcgacgggtca 661
QY 841 TGAGTGTGTGTGGGGAGAGCTATGAGTGGCTCTGAGAGGATGAGATCTCTCT 900
DB 662 tgagtggtgttgaggagagcttatgagtggtctgagagagagagagagagagagag 721
QY 901 AACCGGCCCCCTGTGGGAGAGCAATGATTTATCTGAGGCCGTGACCTCAAGATCA 960
DB 722 aacggccccctgtgggagagacaatgatatttatcgaagccgtgactcaagagctca 781
QY 961 CCTATGAGGAGACATACCTGGGAGGAGTGAATGAGGAGTGGGAGGAGTGGGAGATG 1020
DB 782 cctatgagagacatacctgtggagagctgacgtatgggggtctggccagctggagatg 841

QY 1021 GTGTGGTGGGCTGCATGACTTTAGAGAGATGTCAGAGAGCTGGGGTGGCCAGGCTATG 1080
DB 842 gtgtgggtggctgcatgactttagaagagatgcagagagctgggtctggccagctatg 901
QY 1081 ACTATGAGGAGAGGAGCAACACAGCTTCTCCAGTGGCTATGATGATGATGATGATGAT 1140
DB 902 actatgagagagagagcaacacagcttctccagtggtatgtgagatgtgattgattg 961
QY 1141 TTGACCGGCTGAGGAGCTTCCAGGCTATGAGTGGCTGATCAACAATGCAACAGCTGG 1200
DB 962 ttgaccggctgagagagcttccagagctatgagagtgctcactgtacaacaatgcaacagctgg 1021
QY 1201 GAGCCGCTGCTGCTGGGAGGAGTGAATGCTGCTGGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1022 gagccgctgctggcggggtggaatgtgcttccggcggtgcccctgcagagcctggg 1081
QY 1261 AGGGAGACCCATGAGGAGCAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
DB 1082 agggagaccatgagagcaacacagagggagacactggggagaccagagccgggctg 1141
QY 1321 TCTGAGTCCCTTGGGAGGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 1142 tctgagtgcccttggcggcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1201
QY 1381 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1202 ggcctgtgtactcttccaggaatctcctcactcactcactcactcactcactcactc 1261
QY 1441 CGGACATGGGAGGAGCCTTCCGCGAGCCCTGATGGGAGGAGGAGGAGGAGGAGGAGG 1500
DB 1262 cggacatgggagagccttccgagccctggtgtgtgtgtgtgtgtgtgtgtgtgtgt 1321
QY 1501 ACTTACAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1557
DB 1322 acttacagagtggtggagtggtggagtggtggagtggtggagtggtggagtggtggag 1381
QY 1558 GAGACCCGAGCCGATCTGATGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1617
DB 1382 gagaccgagcccatctcactcactcactcactcactcactcactcactcactcactc 1441
QY 1618 TCAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1677
DB 1442 tcaatgctgatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1501
QY 1678 GGGGTTGGAGAGGAGCTACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1737
DB 1502 ggggttggagagagagcttgcagcttgcagcttgcagcttgcagcttgcagcttgc 1561
QY 1738 ACAACGCGCAGAGTCTAGAGAGCAACCCCGTACAGAGAGGAGGAGGAGGAGGAGG 1797
DB 1562 acaacgcgccagagcttagagagcaaccccgtaacagagagagagagagagagagag 1621
QY 1798 CGGCCACTCCGCTCCCTGTGTCCTCAATGAGTCTGCTGCTGCTGCTGCTGCTGCTG 1857
DB 1622 cggccactccgctccctgtgtcccaatgagctgtgtgtgtgtgtgtgtgtgtgtgt 1681
QY 1858 ACCGCTCTCTTCTGGGAGCTTACGCGCTTCCCTGAGAGCCCGGAGCCCGCACACCG 1917
DB 1682 accgctctcttctgggagcttaccgcttccctgagagagagagagagagagagag 1741
QY 1918 CTTGGGCAAAACCAACCAACAGGCTTACAGTGGGAGGAGGAGGAGGAGGAGGAGG 1977
DB 1742 cttgggcaaaaccaaaccacacacagagcttaccagagagagagagagagagagag 1801
QY 1978 CAGGGGCCCCGCTTGTGCCCCACCTCCACAGAGAGCTCCCAATTATGCGGAGGCTG 2037
DB 1802 caggggccccgcttgtgccccactccacagagagagagagagagagagagagagag 1861
QY 2038 ACATGTTAACCTGAGGAGGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2097
DB 1862 acatgttaacctgagagagctgacaggggagagagagagagagagagagagagag 1921
QY 2098 CAGGGGAGTGGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2157


```
Db 1423 ctagcgcggggccacacatagctacacaaatgctgcgcatggtgcgagcccaagctccct 1482
Oy 2593 CCGGATGACCTATCTGAGCAGACACTAACTTTGATCGGAGACCTGGACGCGGAGACT 2652
Db 1483 ccggacatgcctatctcggccacacacacttctacatcggagcctcggcagcgagact 1542
Oy 2653 GCGTAGTGGGGAAAATTTCACATCAAAATCGCAGACTTTGGCAGTACGCCGGAACCTCT 2712
Db 1543 gccatgttgggaaaatttcacacatcaaaatcgagacttggcatgagccggaaacctc 1602
Oy 2713 ATGTGGGGACTATTACCGTGTGAGGGCGGAGAGTGCCTCCATCCGCTGGATGGCCT 2772
Db 1603 atgcctgggactattacacgctcgaagcgccggagctgcacatccgctggaatgacct 1662
Oy 2773 GGGAGTCATCCCTCATGAGGGGAAGTTACGACTGCGAGTGAAGTGGGCCCTTGGTGTGA 2832
Db 1663 gggagtcacatccatcctcgaagttcgaagtcgaagtcgaagtcgagcttgcgtga 1722
Oy 2833 CCTGTGGAGTGTCTGATGCTGTGTAGGGCCAGCCCTTTGGGACGCTCACGACGAGC 2892
Db 1723 cccgtgaggagtgctgtagtctcgtatggccagcccttggcagctcaacgagagc 1782
Oy 2893 AGGTATCGAGAACGCGGGGAGTTCTTCGGGACAGGGCGGAGGTGACTGTCC 2952
Db 1783 agtcatcagaaacgagcgaggagttcttcggacagcgcgagctgactgctgcc 1842
Oy 2953 GCGCGCTGCTGCTGCGCCGAGGCGCTATATGAGTGATGCTGCTGGTGGAGCGCGGAGT 3012
Db 1843 ggcgcgctcgtcgtccgca .ggcyataatgagctgagctcgtgctgagcgagcgaggt 1901
Oy 3013 CTGAGCAGCGACACCCCTTTCCAGCTGCATCGGTTCTTGCGAGAGGATGACTACACA 3072
Db 1902 ctgagcagcgacacaccttcccaagctgcatggttccctgcgagagatgactcaaca 1961
Oy 3073 CGGTGTGAATACACATCCAGCTGCCCTCCCTCAGGAGTGAATCCAGGGAGCCAGTG 3132
Db 1962 cgggtgatacacatccatccagctgccctccctcagagagatccagggagcaggt 2021
Oy 3133 ACATTAACAGAGAGACACAATGGACACTGCGCTTCCCTCCGACACCATCAC 3192
Db 2022 aactataaacaagagaaatgacacatgacacttgccttccctcccgacacacacac 2081
Oy 3193 TCTATAGAGCAGTGAAGTGAAGTGGAGTGGAGCCACCGAGAGTATGCGCCCT 3252
Db 2082 tctatagagcagtgagacgcagcggtggtggtcccaacccagagagctgagctccct 2141
Oy 3253 CTCCCTTCTGAGACACTCTCATGTCCTTCCCTGTTCTTCTCTAGAACCCCTG 3312
Db 2142 ctcccttctcgagacacatctcatgctcccttctcttctctctagagccctg 2201
Oy 3313 TCGCCACCCAGCTGCTGTGTGATGGATGCGATCTCCACCTCTCTACGATCCCTG 3372
Db 2202 tcgcccacacagctggtctctgtgagtggatctctccacctctctcagcatccctg 2261
Oy 3373 GGGAAAGGTGGGGAATATATAGATAGACACTGACATGGCCCATTTGAGACCTGGGC 3432
Db 2262 gggaaagggtgggaaataatagatagacatgagacatggcccatggagcacttgggc 2321
Oy 3433 CCCACTGGAACAACACTGATTCCTGAGAGAGTGGCTGCG -CCCCAGCTTCTCTCCCTGT 3491
Db 2322 cccactggaacaactgattctcctgaggtggtgcgcccacagcttctctcctcgt 2381
Oy 3492 CACACACAGGACCCACACAGCTGCTGAGAAATCTGGGGTGAAGAGACAAAGAGAGGAAA 3551
Db 2382 cacacactggacccacacagctgagaaatctgggggtgagagaaagaagagggaaa 2441
Oy 3552 ATGTTCTTGTGCTGCTCTGTACTGTCTGCTCAGTGGGCTTCTCTCTCTCATCA 3611
Db 2442 atgttctcttgctgctctcgtactctgctctcagcttggcttctctctccatca 2501
Oy 3612 CCTGAACACTGACCTTGGGGGTAGCCCCGCCACGCCCTCATGTAACCCCACTTCCAC 3671
```

```
Db 2502 cctgaacactgacacctggggtagtcgcccgcccccagccctcagtcacccccactccac 2561
Oy 3672 TTGCAGTCTTGTAGCTAGAACTTCTTAAGCTATACGTTTCTGTGAGATAAATATGGG 3731
Db 2562 ttgcagcttctgtagaactctctcagcctctaaagcttctgtcgtgagtaaatattggg 2621
Oy 3732 ATTGGGGGAAAGAGGAGCAAGGCCCATATGCTTGGGGTGGACATCTAGTAGC 3791
Db 2622 attgggggaaagagagagcaagccatagccttgggttggacatctcagtgtagc 2681
Oy 3792 TGCCACATTAATTTCTAATATCACTTGGGGTTGTACATTTTGGGGGAGAGACACA 3851
Db 2682 tgcacatatttctctataatcacttgggttgcacatttgggggagagacaca 2741
Oy 3852 GATTTTACACTAATATATGACCTAGCTTAGGCAATTTAATCCCTGCTAGAGCAG 3911
Db 2742 gattttacactaatatagctagcttgaagcaatttcaatccctcagtagagag 2801
Oy 3912 GTAAATAATAAGTTGAGTTTCCACAAAAAATAAAAAA 3953
Db 2802 gtaataataaggttgaagtttccacaaaaaataaaaaa 2843

RESULT 6
AAH3198
ID AAH3198 standard; cDNA: 2861 BP.
XX
AC AAH3198:
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:254.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 6; ss.
XX
OS Homo sapiens.
XX
PN M020012920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI: 2001-235357/24.
DR P-PSDB: MAG73767.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX
PS useful for preventing, diagnosing and/or treating colorectal cancers -
XX
Claim 1; Page 2388-2389; 9803pp; English.
XX
AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
```

CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC

50 Sequence 2861 BP; 567 A; 899 C; 800 G; 589 T; 6 other;

Query Match	Score 2783.6;	DB 22;	Length 2861;
Best Local Similarity	99.08;	Pred. No. 0;	
Matches 2833; Conservative	5;	Mismatches 4;	Indels 20; Gaps 3

[illegible]

QY	1993	TCGCCCACTCCCAAGAACAGCTCCCCATTATGCCGAGCTGACATTGTACCTTC	2052
Db	901	TCGCCCACTCCCAAGAACAGCTCCCCATTATGCCGAGCTGACATTGTACCTTC	960
QY	2053	AGGGGATACGGGGGGGAAACACCTATGCTGTGGCTTGACATGCCCAAGGGCACTGGGG	2112
Db	961	AGGGGATACGGGGGGGAAACACCTATGCTGTGGCTTGACATGCCCAAGGGCACTGGGG	1020
QY	2113	ATGGGCCCCCAGAGTGTGATTTCCCTCATCTCCGATCCGGCTTCAAGGAGAAGCTGGCG	2127
Db	1021	ATGGGCCCCCAGAGTGTGATTTCCCTCATCTCCGATCCGGCTTCAAGGAGAAGCTGGCG	1080
QY	2173	AGGGCAGATTMTGGGAGGTGCACCTGTGTGAGTGCAGACGCCCTCAAGATGTGGTCACTC	2232
Db	1081	AGGGCAGATTMTGGGAGGTGCACCTGTGTGAGTGCAGACGCCCTCAAGATGTGGTCACTC	1140
QY	2223	TTTGATTTTCCCCCTTAATGTCGTGAAGGAGACACCTTTGCTGTAGCTGTCAAGATCTTAC	2292
Db	1141	TTTGATTTTCCCCCTTAATGTCGTGAAGGAGACACCTTTGCTGTAGCTGTCAAGATCTTAC	1200
QY	2293	GGCCAGATGCCACCAAGATATCCAGCTTCCTCTGTCTCCAGGAATGATTTCTGTAAAG	2352
Db	1201	GGCCAGATGCCACCAAGATATCCAGCTTCCTCTGTCTCCAGGAATGATTTCTGTAAAG	1242
QY	2353	AGGTGAAGATCATGTGAGGCTCAAGAGACCCCAATCATTTCCGCTCTGTGGGCTGTGTG	2412
Db	1243	AGGTGAAGATCATGTGAGGCTCAAGAGACCCCAATCATTTCCGCTCTGTGGGCTGTGTG	1302
QY	2413	TTCAGGACACGACCCCTCTGTCATGATTTACTGACTACATGAGAGAACGGGACCTCAACAGT	2472
Db	1303	TTCAGGACACGACCCCTCTGTCATGATTTACTGACTACATGAGAGAACGGGACCTCAACAGT	1362
QY	2473	TCCCAATGTCCACACACACTGTGGAGAGAACGAGCGAGGGGCCCCCTGGGGAGCGGGCAGG	2532
Db	1363	TCCCAATGTCCACACACACTGTGGAGAGAACGAGCGAGGGGCCCCCTGGGGAGCGGGCAGG	1422
QY	2533	CTGGCAGAGGGGCCACCATCAGCTACCAATGTGCTGTGCAATGTGGAGCCACAGATCGCCT	2592
Db	1423	CTGGCAGAGGGGCCACCATCAGCTACCAATGTGCTGTGCAATGTGGAGCCACAGATCGCCT	1482
QY	2593	CCGGCAGAGGCCCTATCTGTGGCCACACTCACTTTGTATCGGAGACCTGGCCACCGGAGACT	2652
Db	1483	CCGGCAGAGGCCCTATCTGTGGCCACACTCACTTTGTATCGGAGACCTGGCCACCGGAGACT	1542
QY	2653	GCCATGTTGGGGAAATTTTACCATCAATAATCCAGACCTTTGGCAGATGACCCGGACACTCT	2712
Db	1543	GCCATGTTGGGGAAATTTTACCATCAATAATCCAGACCTTTGGCAGATGACCCGGACACTCT	1602
QY	2713	ATGCTGTGGGACATTTTACCGTGTGCAGAGGCCGGGACAGTGCCTCCATCCGCTGTGATGGCT	2772
Db	1603	ATGCTGTGGGACATTTTACCGTGTGTGCAGAGGCCGGGACAGTGCCTCCATCCGCTGTGATGGCT	1662
QY	2773	GGGAGTGCATTCATGATGGGAAATTTTACAGACTCGAGTACAGTGTGGGCTTTGTGTGTGA	2832
Db	1663	GGGAGTGCATTCATGATGGGAAATTTTACAGACTCGAGTACAGTGTGGGCTTTGTGTGTGA	1722
QY	2833	CCCTGTGGGAGTGTGATCTCTGTAGGGCCCAAGCCCTTTGGGCAAGCTCAACGAGAGAC	2892
Db	1723	CCCTGTGGGAGTGTGATCTCTGTAGGGCCCAAGCCCTTTGGGCAAGCTCAACGAGAGAC	1782
QY	2893	AGGTATCGAGAAACGGGGGGAGTTCTTCCGGGACCAAGGCCCGGACAGTATCTCTCC	2952
Db	1783	AGGTATCGAGAAACGGGGGGAGTTCTTCCGGGACCAAGGCCCGGACAGTATCTCTCC	1842
QY	2953	GGGCGGCTGCGTCCCGGAGAGGCGCTATATGAGCTATGCTTGGTGTGCTGGAGCCGGGAGT	3012
Db	1843	GGGCGGCTGCGTCCCGGAGAGGCGCTATATGAGCTATGCTTGGTGTGCTGGAGCCGGGAGT	1901
QY	3013	CTGAGCAGGACACCCCTTTCCAGACCTCATCCGGTTCCTGGCAGAGATGACATCAACA	3072
Db	1902	CTGAGCAGGACACCCCTTTCCAGACCTCATCCGGTTCCTGGCAGAGATGACATCAACA	1961
QY	3073	CGGTGTGAATCACACATCCAGCTGCCCTCCCTTCAGGAGTATCCAGGGCAAGCCAGAG	3132

```

Db 1962 cggcttgatcaccatcagctgcctccctccctcagggagatccagggagagccatg 2021
Oy 3133 ACACATAAACAAGAGACACAATGGACACTGCTGCCCTTCCCTCCCGACAGCCATCAC 3192
Db 2022 aacataaacaagagagacacatgacactcctgcctccctccctcccgagacacacac 2081
Oy 3193 TCTAATAGAGCAGTGAAGATGACAGTGGGCTGGGCGCCACCCAGGAGCTGATGCCCTT 3252
Db 2082 tctaatagagcagtgagactcagctgagctgagctgagctgagctgagctgagctg 2141
Oy 3253 CTCCCTCTCTGAGACACTGCTGATGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3312
Db 2142 ctccctctctgagacacactcctcctcctcctcctcctcctcctcctcctcctcct 2201
Oy 3313 TCGCCACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3372
Db 2202 tcgccacacagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2261
Oy 3373 GGGAAAGGTGGGGGAAATATAGATAGACACTGACATGGCCCAATGGAGCCTGGGC 3432
Db 2262 gggaaagggtgggggaaatataagataagataagataagataagataagataagata 2321
Oy 3433 CCCACTGACACACTGATCTCTCTGAGAGTGGCTGGCG-CCCCACTCTCTCTCTCTCTCT 3491
Db 2322 cccactgacacactgactcctcctgagagctgctgagagctgctgagagctcctcctc 2381
Oy 3492 CACACATCGACCCACATGCTGAGATCTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAA 3551
Db 2382 cacacatcgacccacatgctgagatcctgagagctgagagagagagagagagagagaa 2441
Oy 3552 ATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3611
Db 2442 atgttctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2501
Oy 3612 CCTGAAACACATGAGCTGGGGGTAGCCCGCCCGCCCGCTGATACCCCGCTCCCTCCAC 3671
Db 2502 cctgaaacacatgagctgggggtagcccgcccgcccgccctcctcctcctcctcctc 2561
Oy 3672 TTGCACTCTGAGCTAGAACTCTCTAAGCTTACCTTCTGAGTAAATATTTGGG 3731
Db 2562 ttgcagctctgagctagaaactctcctcctcctcctcctcctcctcctcctcctcct 2621
Oy 3732 ATGGGGGGGAAAGAGAGACAAAGCCCAATGCTGGGGGTGAGCATCTAGTGTAC 3791
Db 2622 atgggggggaaagagagagagagagagagagagagagagagagagagagagagag 2681
Oy 3792 TGCCACATTTGATTTCTATATCACTGGGGTGTGATCAATTTTGGGGGAGAGAC 3851
Db 2682 tgccacatgatttctcctacacactcctgagagagagagagagagagagagagagag 2741
Oy 3852 GATTTTACATAATATATGAGACTGAGCTGAGCAATTTTAACTCCCTGCACTAGGAG 3911
Db 2742 gattttaccataataatgagagagagagagagagagagagagagagagagagagagag 2801
Oy 3912 GTAATATATTAAGGTGAGTTTCCACAAAATAAAAAA 3953
Db 2802 gtaataataaaggttgcgttctccacaaaaaataaaaaa 2843

```

RESULT 7

AAV48292 ID AAV48292 standard; cDNA; 3096 BP.

AAV48292;

16-NOV-1998 (first entry)

DiscoIdin domain receptor 2 gene.

DiscoIdin domain receptor; transformation; metastasis; collagen; MMP-1.

Cleidocranial dysplasia; Stickler syndrome; extracellular matrix; MMP-1.

```

OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 354..2921
FT sig_peptide 354..416
FT mat_peptide 417..2918
FT /tag= a
FT /product= "DiscoIdin domain receptor"
FT /tag= b
FT /tag= c
FT /tag= c
PN W09834954-A2.
PD 13-AUG-1998.
PE 05-FEB-1998; 98MO-CA00093.
PR 06-FEB-1997; 97US-0041578.
PA (MOUN ) MOUNT SINAI HOSPITAL CORP.
PI Pawson A, Vogel W;
DR WPI. 1998-447168/38.
DR P-PSDB; AAM77114.
XX
XX Novel ligands of discoIdin domain receptor tyrosine kinase,
XX especially collagen - useful for treating e.g. metastasis,
XX cleidocranial dysplasia or Stickler syndrome
XX
XX Disclosure: Fig 22a; 115pp; English.
XX
XX The DDR can be used to identify and evaluate substances which affect DDR
XX receptor tyrosine kinase signalling pathways in the cell. Compounds
XX which modulate such signalling pathways can be used to alter
XX transformation or metastasis in mammals. To treat conditions involving
XX structural or functional deregulation of collagens, e.g. Cleidocranial
XX dysplasia or Stickler syndrome, conditions requiring modulation of
XX extracellular matrix synthesis, degradation or remodelling, or to treat
XX conditions needing modulation of MMP-1 expression such as wound healing.
XX
XX Sequence 3096 BP; 762 A; 791 C; 752 G; 791 T; 0 other;
SQ
Query Match 16.2%; Score 642; DB 19; Length 3096;
Best Local Similarity 56.5%; Pred. No. 5.3e-131;
Matches 1534; Conservative 0; Mismatches 975; Indels 207; Gaps 10;
Oy 349 TCGTCTGCTCTGTTGGCAGTGGAGATGCTGACATGAAAGGACATTTGATCTGCA 408
Db 379 Tgtgtctgtctctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtct 438
Oy 409 AGTGGCGATGCTCCCGGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
Db 439 tatgcgcatacctcctgagagagagagagagagagagagagagagagagagagagagag 498
Oy 469 CCAGCTCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
Db 499 ccagtcagtggtcagagagagagagagagagagagagagagagagagagagagagag 558
Oy 529 ATGGGGCTGTGGCCCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 585
Db 559 atggagcctgtgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 618
Oy 586 TGGATCTACAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
Db 619 ttgacttgacacccccaattatcactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 678
Oy 646 GCGTGGCAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
Db 679 gtcatgagcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 738
Oy 706 GGATGGCTGGAAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765

```


Db 739 ggaatcctcttggcgaaccgcatgagaaacagtgctgataagtaagtaacccctatg 798
Qy 766 gagggtgctcaaaagaccttggcccccacatggtggccgacctggtgcttcaacccc 825
Db 799 acaatttcctaaagagacttggagccgacattgtagccagatttgcggttcatccag 858
Qy 826 gggctgaacgggctgatgctgctgctgctgctgctgctgctgctgctgctgctgctg 885
Db 859 tcacaggaacacatccatgatactgctgataagtgagagcttcaagcgtgctgctgctg 918
Qy 886 atgagacttctgcttcttaccgccccctgctgctgctgctgctgctgctgctgctgctg 939
Db 919 atggccttggtgcttcaacagctccagcgtgagcagcttgccttccctgaggttcca 978
Qy 940 cccgcttaccttcaaacgacttccacctaagacacataccgtggcgacctgacgtatggcg 999
Db 979 tcaattatctgaatgattctgctatgataagctgctgatacagcaagcaagaagg- 1037
Qy 1000 gctgagccacacctgacagatgctgctgctgctgctgctgctgctgctgctgctgctg 1059
Db 1038 --caggcccaattgacagagtgctgctgctgctgctgctgctgctgctgctgctgctg 1095
Qy 1060 tgcgggcttgcgacgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1119
Db 1096 accacgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1155
Qy 1120 atgctgagatgagacttggatgctgctgctgctgctgctgctgctgctgctgctgctg 1179
Db 1156 acattgagatgatttggatgctgctgctgctgctgctgctgctgctgctgctgctg 1215
Qy 1180 gtaacacatgacacacacctgagagccctgctgctgctgctgctgctgctgctgctgctg 1239
Db 1216 gcaacaacatgatttggataagagcttgaagagcttgaagagcttgaagagcttgaag 1275
Qy 1240 gtagcccttgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1299
Db 1276 ctg---aagcagtgagtgaggaacctaattgcaatttcccttcccttgcctgataag 1332
Qy 1300 gggaccccgacagcccgctgctgctgctgctgctgctgctgctgctgctgctgctg 1359
Db 1333 tcaaccacagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1392
Qy 1360 agtgcgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1419
Db 1393 agtgcgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1452
Qy 1420 atgagtgagacaaatcttcttcttcttcttcttcttcttcttcttcttcttcttct 1479
Db 1453 atgctgcaaatgtaacaacacttgaagccctgcccacaccttccca----- 1497
Qy 1480 cggctgcccac 1539
Db 1498 -----tgac 1518
Qy 1540 cccctgcccac 1599
Db 1519 caatgcttaaatgtaagac 1578
Qy 1600 tcttgccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1659
Db 1579 tcttgccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1638
Qy 1660 tcagacagctgagacagagcttggagagagagagagagagagagagagagagagagag 1719
Db 1639 tggag 1688
Qy 1720 gggacacatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1779
Db 1699 gtgattcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1731
Qy 1780 cccggccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1839
Db 1732 catcaactagtgatacaagaggttccaaactctgacttaagatcgcattctt----- 1778
Qy 1840 tgccttccaaac 1899
Db 1779 -----cccttgcgc 1788
Qy 1900 cgggccccccac 1959
Db 1789 ctgactacacagagac 1848
Qy 1960 atatgagacctgagaaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2019
Db 1849 agagatcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1908
Qy 2020 cccattatgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2079
Db 1909 cccactatgcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1968
Qy 2080 ctgtccctgcaactgccccccac 2136
Db 1969 cagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2028
Qy 2137 ctgcattcgcagctcgccttcaagagagagagagagagagagagagagagagagagag 2196
Db 2029 ccaggaac 2088
Qy 2197 tctgtgagctgac 2256
Db 2089 tctgtgag 2148
Qy 2257 agggac 2316
Db 2149 ccaacacagcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2205
Qy 2317 gcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2376
Db 2206 -----ccagagatgatttcttcttcttcttcttcttcttcttcttcttcttct 2250
Qy 2377 aggacccac 2436
Db 2251 aggacccac 2310
Qy 2437 ttactgacttactgag 2496
Db 2311 tcaactgacttactgag 2370
Qy 2497 acaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2556
Db 2371 attcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2403
Qy 2557 acccaatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2616
Db 2404 acaccaatcgaagcttacttacttacttacttacttacttacttacttacttacttacttact 2463
Qy 2617 tcaacttggtaac 2676
Db 2464 ttaatttggtaac 2533
Qy 2677 tcaaaatgcaagacttggag 2736
Db 2524 tcaagatagctgacttggag 2583
Qy 2737 agggcggcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2796
Db 2584 agggcggcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2643
Qy 2797 tcacagctgag 2856
Db 2644 tcaactagcaag 2703
Qy 2857 gtaagcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2916
Db 2704 gtcaagagacagccctattccacagctgacagatgaaacaggttattggaatacagagag 2763

OY	1420	AGTGTGAACAATTCTCTCCGGACTTGGAGGACCTTCCCGCAACCCCTGTGGC	1479
Db	1469	atgtctgaatgaacaacatcttgaagcccttgcacactctctca-----	1513
OY	1480	CGCCTGGCCACCTCCACCACCAATTACAGACTTGGAGCTGGAGGCCGAGCCAGCAGC	1539
Db	1514	-----tggacccaacacatgtatc	1534
OY	1540	CCGTGGCCAGGCGGAGGGAGCCGACCGCATCTCATCGCTGCTGTGGCCATCA	1599
Db	1535	caatgcttaaaagttgatgatcagacaacactcgatcccgattgctgtctgtgtgcataca	1594
OY	1600	TTCGCTCCCTCTGCTCATATTTCCTCATGCTCTGTGGCGGCTGACGTGGCGAGGCTCC	1659
Db	1595	ttcttatccctcttgcacatcatcttctcatcatcctctcttgaagtcagttctctgagaanaatgc	1654
OY	1660	TCACCAAGGCGTGAAGGAGGCTGTGGAAAGGAGGCGTCAAGCTTCCTCTCTCTG	1719
Db	1655	tggagaaaggtctcttcggaagatgctctgataatgaaagaaagtaagcttcccttgcacaa	1714
OY	1720	GGGAGCTATCTCATACAGACCGCCAGGCTCTTAGAGAGCCACCCCGTAGAGAGC	1779
Db	1715	gtgattctagatgttcaacaat-----aacgcctct	1747
OY	1780	CCCGGCTCTGTGGGAATCCGCCCATCTCGCTCTCTGTGCCAATAGCTCTGCTTGC	1839
Db	1748	catcaactaagtgaacaagggtccaaactcgacttaagtcagtcgcatctt-----	1794
OY	1840	TGCTCTCAATCCAGCTTACCGGCTCTTGTGGCACTTACGCCCGTCCCTCTGAGGCC	1899
Db	1795	-----ccctctggc	1804
OY	1900	CGGCGCCGCCACACCGCCTGTGGCCAAACCAACACCCAGGCTTACATGAGGGAGCT	1959
Db	1805	ctgactaacagagcatccacagctgataagaaactccagaatttgcacaggggaag	1864
OY	1960	ATTGTGAGCTTGAGAAAGCCAGGCGCCCGCTTGTGCCCCACCTCCCGAGAACAGCTCC	2019
Db	1865	aggaagtcaggtctgcagcgggtgtgtgtgaagccagtcacagccagtcgacctgaggggtgc	1924
OY	2020	CCGATTATCCGAGGCTGACATGTGTATCCCTGTGAGGCGTCACCGGGGCAACACTATG	2079
Db	1925	cccacatgtcagaggtctgcataatgtbaacctccaaagagtgaaaggaacaaactact	1984
OY	2080	CTGTGCTTGCACCTGCCCGCAGGGGCGCAGTC---GGGATGAGGCCCCCAAGTGGATTTC	2136
Db	1985	cagtcgctgcgtcaacacatgcagccctgctctcagaaaaaagtgtgtgtgagggagtcc	2044
OY	2137	CTCGATCTGACATCCGCTTCAAGGAGAGACCTTGGCGAGGCGCATTTTGGAGGTGCACC	2196
Db	2045	caaggaaaactctaacttcaaaagaaagcttgagaaagacagtttggggagttcatc	2104
OY	2197	TGTGTAGGTGCACAGCCCTCAAGATCTGTGTGATGCTTGAATTTCCCTTATATGTGCTGA	2256
Db	2105	tctgtgaagtgtgaaggaaatgagaanaattcaagacaaatattgcacctgaagtgcagtg	2164
OY	2257	AGGGAACCCCTTGTGCTGTAGCTGTCAAGATTTAAGGCGCAGATGCCACCAAGAAATGCCA	2316
Db	2165	ccaaccagcctgtctgtgtgtgtgcgaanaatgctccgagcagatgccaacaaagaatg---	2221
OY	2317	GCTTCTCTGTTTCTCCAGAGATGATTTTCTGAAAGAGGTGAAGATGATGTGAGGCTCA	2376
Db	2222	-----cgaagaaatgtatttcttaagaaataaaagatacatgtctgcgtca	2266
OY	2377	AGGACCCCAACATCATTTGGGCTGTGGGCGCTGTGTGTGCAGAGAGACCCCTGTGACATGA	2436
Db	2267	aggacccaacatcatcatctatctatctgtgtatatactatgaaacccctctgtataga	2326
OY	2437	TTACTGACTACATGGAGAAGCGGACCTCAACACGATTCTCATGTCGCCACACCGAGGAGG	2496
Db	2337	tcaactgaatacatggaagaatgagatctcaatcagttcttcttcccgcaagagccctca	2386
OY	2497	ACAAGGCAAGCCGAGGGGCCCCCTGGGAGCGGGCAGAGCTTGGCAGAGGCCCAACATCAAGCT	2556

OY	2080	CAGGCCGTGCACCTGCCCCCAGGGCCAGTC---GGGATGGCCCCCCCAGATGGATTTC	2130
Db	1985	cagtgctgcgcgtacatcatgtagcctgtctctcaagaagaatgtygtctgtgaagagtctcc	2044
OY	2137	CTCGATCTGCACCTCCGCTTAAAGAGAAGCTTGCGAGGGCCACTTTGGAGGGAGCC	2196
Db	2045	ccaggaaactcctaacttccaagaagagactggyagaagacagtttggysaagtttcac	2104
OY	2197	TGTGTAGGTGCACAGCCCTCAAGATCTGGTCACTGTGATTTCGCCCTTAATGGCGTA	2256
Db	2105	tctgtgaagctgysagyalgaaaattcaaagaacaagatttgcctcagatgtcagtg	2164
OY	2257	AGGGACACCCCTTTGGCTGAGTGTCTCAAGATCTTAGGCCAAGATGCCAACGAATGCCA	2316
Db	2165	ccaaccgacctctccctggtggtcggtggaataatgctccagacgaatgccaacaagaaty---	2231
OY	2317	GCTTCTCTTGTCTCCAGGAATGATTTCCCTGGAAGAGGGTGAATAATCATGTCCGAGGCCA	2376
Db	2222	-----ccaggaatgatlltctttaagsygaataaagaatcatgtctcgcctca	2266
OY	2377	AGGACCCCAACATCATTTGGGCTCTGGGCGTGTGTGTGCAGAGAGACCCCTCTGCATGA	2436
Db	2267	aggaccacaacatcacatcatactatctatctgtgtatcatcatgtacacctctctgtatga	2326
OY	2437	TTTTCTGACTCATGAGGAAGAGGGAGACCTCAACCAGTTCCCTCAGTGGCCACAGCTGAGG	2496
Db	2327	tcactgaatacaatgysagaatgysagatctcacaatgattcttcccgcacagagcccccta	2386
OY	2497	ACAAGGACAGCCGAGGGGGCCCCCTGGGGACGGGACAGCTGGCGAGGGGCCCACTATCACGT	2556
Db	2387	attcttcctccagcg-----atgtacgcactgtcagtg	2419
OY	2557	ACCCATATGCTGCTGCAATGTGTGGCAGCCAGATTCGCTCTCCGCAATCGCTATCTGGCCACAC	2616
Db	2420	aaccacatctgaagtgattatgtctaccaccaattgcctctgycatgaaagtaccttccctc	2479
OY	2617	TCAACTATGTACATTCGGAGACTGGCCAGCGGAGAACTGCCATGTTGGGAAAATTTQACCA	2676
Db	2480	ttaatttgttccacgagatctgycacacagaaactggttaagtgglaaagaactaacaa	2539
OY	2677	TCAAAAATCGACACTTTGGCATGAGCGGAGAACCTCTATGCTGGGAGACTATTACCGTGTGC	2736
Db	2540	tcaagaatacgtgacttgtgaatgagysagactgtacagtggttacttatcaocgctcc	2599
OY	2737	AGGGCCGGGACAGTGTGCCCATTCGCCGTGATGTGCCCTTGGGAGTGCATCTCATGGGAGAT	2796
Db	2600	agggccggsgcagtgctcccatcgcgtlgtatgtcttggysagatlatctltctgyscagt	2659
OY	2797	TGACGACTCGATGTGACGTGTGGGCGCTTTGATGTGACCCGTGGGAGGTGCTGATGCTCT	2856
Db	2660	tcaactacagcagtgatgtgtgtyggcccttggysgttactctgtgysgaacttcaaccttl	2719
OY	2857	GTAGGGCCAGCCCTTTGGGACAGCTCACCGACGAGCAGGATCATCGAAMACGGGGAGT	2916
Db	2720	gtcaagaagaagccctatccacagctgtcagatgaacagattatggaataatcysagagt	2779
OY	2917	TCCTCCGGAGACAGGGCCGGGAGGTATTACTGTGCCGCCGCGCTGCTGCCCGCAGGGCC	2976
Db	2780	ctctccgaaaccaagaagysagaaacttacctccctcaacaagacattgtctcgtgactctg	2839
OY	2977	TATATAGAGCTGATGCTTCCGTGCTGTGGAGCCGGGAGTCTGTGACAGCACCTTTTCCC	3036
Db	2840	tgtataagctgagtctccagctgtcgtgagaagatatcagaaacgcttccctaatccaag	2899
OY	3037	AGCTGCATCGGTCTCT 3052	
Db	2900	aaatccaccttgcgt 2915	

RESULT	10
AAT93784	
ID	AAT93784 standard; cDNA; 3157 BP.
XX	

AC	AA193784;
XX	
DT	16-FEB-1998 (first entry)
XX	
DE	CCR-2, a human mammary carcinoma kinase 10 MCK-10) family member, cDNA.
XX	
XX	Mammary carcinoma kinase: MCK-10; CCK-2; receptor tyrosine kinase;
KW	proliferative disease; cancer; MCK-10 activity; aberrant expression; ds.
XX	
OS	Homo sapiens.
XX	
FX	
FX	Key Location/Qualifiers
FT	370..2934
FT	CDS
FT	/*tag- a
XX	
XX	
PN	US5677144-A.
XX	
PD	14-OCT-1997.
XX	
PF	08-NOV-1994; 94US-0336343.
XX	
PR	16-NOV-1993; 93US-0153397.
XX	
PA	(ALVE/) ALVES F H E.
PA	(ULTR/) ULLRICH A.
XX	
PI	Alves FHE, Ullrich A;
XX	
DR	WPI: 1997-511869/47.
XX	
DR	P-PSDB: AAM34671.
XX	
PT	Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PT	for it, useful for cancer diagnosis
XX	
PS	Claim 9; Fig 3; 70pp; English.
XX	
CC	
CC	The present sequence represents the cDNA sequence of human CCK-2, a
CC	member of the mammary carcinoma kinase 10 (MCK-10, AA193785) family of
CC	receptor tyrosine kinases. Expression of CCK-2 is associated with
CC	proliferative diseases such as cancer. The CCK-2 gene was identified by
CC	PCR and a cDNA prepared from colonic adenocarcinoma RNA. CCK-2 is
CC	expressed in a wide variety of cancer cell lines and tumour tissue. The
CC	CCK-2 nucleic acids can be used for diagnostic purposes to detect
CC	aberrant expression of CCK-2 genes. Engineered cell lines, containing
CC	recombinant vectors with the present sequence, are useful for producing
CC	infectious retroviral particles. The cell lines may also be used to
CC	evaluate and screen drugs involved in CCK-2 activation and regulation.
XX	
SO	Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T; 0 other;

Query Match	16.2%	Score 662	DB 181	Length 3157
Best Local Similarity	56.5%	Pred. No. 5.3e-131		
Matches 1534	Conservative	0	Mismatches 975	Indels 207
			Gaps	10
QY	TCGCTGCTGCTCTGGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGATCCCGCA	408		
Db	395 tggctgtcttctctgtctgcctatcttgagttctgcgaagctcagyttaatccagca	454		
QY	409 AGTGGCGGTATGCCCTGGGACATGACGAGACCGGACCATCCAGACATGATCTGCTT	468		
Db	455 tatgcgcattatctctctggcattcgaagggcagattccagatggagcatcacagctt	514		
QY	469 CCAAGTCCGTGGGCAATTCACATCCAGTCCGCCGCCACAGCAGGTGGAGAGACATGACGGG	528		
Db	515 ccaagtcagtgvcagagtcacagctgcacaaatatygaagygcttgatccagaagaag99	574		
QY	529 ATGGGCGCTGTGATGCCGCCAGAGGTGGTGTTCCTCCAGGA--GGAGGATGACTTGCAG	585		
Db	575 atggagccttggtgcccttgagatccagtggaacctgatygaactgaaagatcttcgcaga	634		
QY	586 TGGATCTACAAACGATCCACCTGATGCTGTGGTGGGACCCAGGAGCGGATGCCGGGG	645		

QY	2019	CCCCATATGCGCAGGCGTGCATATGTTTACCTGTGACGGCGCCACCGGGGGCAACACCTAT	2078
Db	2039	ccccactatgcaagaagccgcacatagtgaaatcctccaaagagtgtaagaagtggaataacacctac	2098
QY	2079	GCCTGCGCTGCACACTGCCCCAG---GGCAGTGGGGATGGGCCCCCCAGAGTGATTTTC	2135
Db	2099	tgtgtctgcgtctgtaaccacagatctgctatccgggaagaatgtygactgtggaagagtct	2158
QY	2136	CCNCGATCTGAGACTCCGCGTTCAAGGAGAAACCTTGGCCAGGGCCAGTTTGGGAGGTGCAC	2195
Db	2159	ccccagaaactgtcttgcgccttccaaaggaaagcttgggaagaagccaaagtcttgggaagttcat	2218
QY	2196	CTGTGTGAGGTGCAGACAGCCCTCAAGATCTTGCTGATAGTCTTATTTTCCCTTAAATGTGCGT	2255
Db	2219	ctctgtgaaagtggagggaatgtaagaatcaagaacaaagtcttgcactagatgtcaagt	2278
QY	2256	AAGGAGACACCCCTTGGTGGTAGCTGCACAGATTTACGGCCAGATGGCCACCAAGATATGC	2315
Db	2279	gccaaacaaagccttgccttctgttgcgcgtgaaatctctccagacagatgccaacaagaatgc--	2336
QY	2316	AGCTTCTCTGTGTTCCAGAGATGATTTTCTGAAAAGGTGGAAGATCATGTGAGGCTC	2375
Db	2337	-----ccggaaatgtattcttcaaaagagatccaagatcagatcgtccgctc	2380
QY	2376	AAGGACCCCAACATCATTTGGCTGTGGCGGTGTGTGTGCAGAGACGACCCCTCTGCATG	2435
Db	2381	aaggaccacaacatcatccgctctttagctttagctgtgtatcatcatgagacccgccttgatg	2440
QY	2436	AATTAAGCATACATGAGAAAGCGCGACCTCACACCACTTCTCTAGTGCACCAAGCTGGAG	2495
Db	2441	atcaagaaatacaatggaatggaatcttaatacaattcttctctgcacgaagcctctg	2500
QY	2496	GACAAAGCAGCGAGGGGCCCTCTGGGAGCGGCGACAGCTGCGAGGGGCCACCATATGC	2555
Db	2501	a-----gttcttgtctatgtatgtccaagtcagt	2530
QY	2556	TACCCAAATCTCTGCAATGTGGACGCCAGATGCGCTCCGGCATGCGGTATGTGGCCACA	2615
Db	2531	tacgcaaaccttgaagtttattgycaaaccagatgtcctctgtatgaagtaccttcgctc	2590
QY	2616	CTCAACTTGTATCATATCGGAGACCTGGGCGACAGCGGAATGCTGCTGTTGGGAAATTTTCAC	2675
Db	2591	ctcaactgttccacacgaaatctgtgccaacagaaaactgttcttagtgggcaagaattacac	2650
QY	2676	ATCAAAATCGCAGACTTTGGCATGTAGCCGAACTCTATGCTGGGAGCATTTACCTGTGTG	2735
Db	2651	atcaagaatagctgatttgtgcatagacagaacactgttacaagtgtgatactacacgatac	2710
QY	2736	CAGGGCCGGGCGAGTGTGCCCATTCGGCTGTGATATGGCCTGGGAGTGCATCTCATGGGAAAG	2795
Db	2711	caaggccgggagtggtctcccactgcgtgaaatgtaagtccttgcgtgggaaagcactctgtgcgcaaa	2770
QY	2796	TTTCAGAGCTGCGAGTGCAGTGTGGGCTTTGTGTGATACCTGTGGGAGGTGGCGTGAATGCAC	2855
Db	2771	ttcacccagcgcaagtgtgtgtggtgccttgggtgagactctgttgggaagaccttaccctt	2830
QY	2856	TGTAAGGCCCAAGCCTTTTGGGCAAGCTCACCCAGACAGCAGAGTATGAGAAAGCGGGGAG	2915
Db	2831	tgcagagagcagcctcatctccagctgttgcatagtacagatgtacatcgagaaacacttgaagag	2890
QY	2916	TTCTTTCGCGGAGACAGAGCGCGGAGAGGTATACCTGTCCCGGCCGCTGGCCGCGCAGGGC	2975
Db	2891	ttctctcgaagaccaaaggagagcagatcatcatcctccataacacgaccttgcgccgcgtctc	2950
QY	2976	CTATATGAGCTGAGTCTTCGGTGTGAGACCGGAGTGTGAGCAGACGACACCTTTTTC	3035
Db	2951	gtgtataagctgtatctcaagctgtctgtgagaaagaaacaaagcacccgcgcatccttcag	3010
QY	3036	CAGCGCATCGGTTCTCTGGCAGAGGATGCATCTCAACAGGGTGTGAATCA	3084
Db	3011	gaataacacctcgtcttcttcaagaagaagccggagtgtatgtatgataca	3059

OY	2487	CAGCTGGAGGCAAGGACGCCGAGGGGGCCOCCCTGGGGAGCAGGGACAGGTCTCGCAGGGGCC	2548		
Db	1940	gagccccaattcttcctccacgg-----atgtaccg	1972		
OY	2547	ACCATCAGCTACCOCATGCTGTCATGTGCAGCCAGATCGCTCCGGCATGCGCTAT	2606		
Db	1973	actgcagttacaccaaatctgaagttagttagtgcatacccacaattcgcctctgacatgaagtaac	2032		
OY	2607	CTGGCCACTCAACTTTTGACATCGGGACCTTGGCACCGCGGAACCTGCTTAGTTGGGAA	2666		
Dp	2033	cttcctccctctaatttgcticaccgagatctgycacacagaacctgyltiagtygtaag	2092		
OY	2667	AATTTCACATCAAATTCAGACGACTTTGGCATAGACCGGAACCTCATAGCTGGGGCATAT	2726		
Db	2093	aactaaccaatcaagaatagatgtaacttggaaatyagcaagyaacctgttacagtygtgaatat	2152		
OY	2727	TACCGTGTGCAGAGCGCGGAGTGCTGCCATCCGCTGGATGGCTGGAGATGCATCCTC	2768		
Db	2153	taccgagatccaggscggcgagygctcccatocgctgtgatgtcttgggagagtatcttg	2212		
OY	2787	ATGGGGAAGTTCAAGACTGCGAATGACGTGTGGCCTTTGGTG	2829		
Db	2213	ctggcagaagtctcacacgaagtgaigtgtgcttggtgg	2255		
<hr/>					
RESULT 14					
ID	AA064158				
AA064158 standard; DNA; 2128 bp.					
XX	AA064158;				
DT	03-FEB-1995	(first entry)			
DE	Partial coding sequence of tyrosine kinase receptor protein.				
KW	Tyrosine kinase; receptor; proto-oncogene; trk; detection;				
KM	diagnosis; antibody; treatment; tumour; antisense; ss.				
OS	Homo sapiens.				
FH	Key	Location/Qualifiers			
FT	CDS	1..1953			
FT	/tag_*	a			
FT	/product=	Partial sequence of protein tyrosine			
FT		kinase.			
PX	DE4239817-A.				
XX	01-JUN-1994.				
PD	26-NOV-1992.				
PF	92DE-4239817.				
PR	26-NOV-1992.	92DE-4239817.			
PA	(CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.				
PI	Holtfrich U, Ruebsamen-waigmann H, Strebhardt K;				
DR	WPI, 1994-184380/23.				
XX	P-PDB; AANR4089.				
XX	New protein tyrosine kinase and related nucleic acid - vectors,				
PT	transformed cells, etc., useful for diagnosis and treatment of				
PT	tumours				
PS	Claim 4; Page 8; 9pp; German.				
CC	The gene is related to the trk proto-oncogene. Antibodies against				
CC	the encoded polypeptide are useful for diagnosis and for the				
CC	treatment of tumours. The antibodies may also be radiolabelled or				
CC	coupled to a cytotoxin for destruction of cancer cells. Antisense				
CC	nucleic acid can be used to inhibit gene expression.				

XX	Sequence	2128 BP; 532 A; 556 C; 504 G; 536 T; 0 other;
50	Query Match	11.4%; Score 453.4; DB 15; Length 2128;
	Best Local Similarity	55.6%; Pred. No. 8.9e-90;
	Matches 1176; Conservative	0; Mismatches 741; Indels 198; Gaps
QY	941 CGGTACCTCAAGACTCCACCTATGACAGGACATACCGTGGCGGACAGCTGCATATGGCG	10000
DB	12 cattactctgaatgattctctcctcatgatgagctgtcttgatatacagcatgacagaaggg--	69
QY	1001 TCTGGGCGACCTGGCCAGATAGTGTGCTGGGGGCGGTGGATGACTTTAGAGAAAGTCAGACT	10600
DB	70 -ctagagcaattgacccaatggtgtgtctgtcgtcgagcttgacccaatgacccaatgaaata	128
QY	1061 GCGGGTCTGGCCAGAGCTATATGATATGGATGGACCAACACACACTTCTCCAGTGGCTA	11200
DB	129 ccaagctgtggccggcgcatctactcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	188
QY	1121 TGTGGAGATGAGTTGATGATTGACCGGCTGAGGGGCTTCCAGGCTATGACAGTCCACTG	11800
DB	189 catgtagatcatgttttgaatttaccgcgcatcagaatttccatcacatgaagctcactgt	248
QY	1181 TAAACAATGACACAGCTGGGAGCCCGTCTGCTGGCGGGGTGGATGTGCTTCCGGC	12400
DB	249 caaacaataatgtttgtctaagaagtgtgaagaatctttaagaagtagacagtgtacttccgtc	308
QY	1241 TGGCGCTGCATGAGCCGAGGAGGAGGAGGAGCCATGCGCCACACACATGAGGGGCAACTGG	13000
DB	309 tg---aagccagtgagtgaggaaccttaatgacatttcccttcccttctcctgtgaagact	365
QY	1301 GGACCCCGAGACCCGGGAGCTGTCTCACTGTCGCCCTTGGCGGCGCTGTGCTCGCTTTTCA	13600
DB	366 caaccacagtgatcggttgtgtcaagctgtccttcccaacacgaatgagcagtgcatcaaa	425
QY	1361 GTGCGCTTCTCTTTTGGCGGGGCGCTGTGTTACTCTTCAAGCAAAATCTCCTTCATCTGA	14200
DB	426 gtttcaataaccatttgcagatataccctgtatgttcaatgagatcacacttccaatcaga	485
QY	1421 TGTGGTGAACAATTCCGTCGCGGCACTGGGAGGACACTTCCGGGCAAGCCCTCGTGGCC	14800
DB	486 tgtgtcaatgtatacaacaactctgaagccctgtccacacttcc-----	527
QY	1481 GCGTGGCCACACTCCACCAACAATTACAGAGTTGGAGTGGAGGCGCCAGAGCCACAGACC	15400
DB	528 -----tatgagaccacaactatgatcc	551
QY	1541 CGTGCCAGAGGCGGAGGAGGAGCCGACCGCATCCTCATGGGCTGCTGTGGCCATCAT	16000
DB	552 aatgtcttaaatgtatgacagcaacactcgatcgtatgtgtgtgtgtgtgtgtgtgtgtgt	611
QY	1601 CCGTCTCTGCTGCTCATATGCTTATGCTGTGCGGGGCTGACACTGGCGCAGGCTCTT	16600
DB	612 cttatctctctcgtgcatcatatgtatcatctctctcgtgagagcttctgtgcagaataatgtc	671
QY	1661 CAGCAGGCTGAGAGGAGGAGGTGTGGAAAGAGAGTCAAGGCTTCACTCTCTGTCGCTGG	17200
DB	672 gggagaaggtcttcggaaggaatgtctgtatgataatgacatgacagcttccctgtcgaag	731
QY	1721 GGACACTATTCATCATCAACAGCCGCGAGGCTCTAGAGAGCACCCCGGTACAGAGACC	17800
DB	732 tgttcttagcatgttcaacaataacacgctctcatca-----	768
QY	1781 CCGGCTCTGTGGAAATCGGCCCATCTCGCTCTCCGTGTGCCAATGGCTTGGCTTGTCT	18400
DB	769 -----cctagtgaacaaggtgtccaactgactaactacgtcatctt-----	810
QY	1841 GCTCTCAATTCAGAGCTACGGCTCTCTTCTGTGGCACTTAGCCGCGCTCCCTCGAGGCC	19000
DB	811 -----ccctctgtgcc	821
Y	1901 GGGCCCCCCCCACACCGCGCTGGGCCAAACCAACCAACACCGAGGCTACAGATGGGGACTA	1960


```

Db      822  tgactaccagagccatccagcgtgatacgaaactccagaaatttgctccaggaggagga 881
Qy      1961  TATGAGACCTGAGAGAGCGAGCGCCCGCTTCTGCCCCACCTCCCGAAGACGCTCC 2020
Db      882  ggaagtcagcgtcagcgtgtgtgaaagccagtcagcccgagtcgctgaggggtgcc 941
Qy      2021  CCATTATGCGGAGCTGACATGTTTACCTCTGAGGGGCTACCGGGGGCAACACCTATGC 2080
Db      942  ccactaagcagaggtctacatagtaaacctccaaagagatgagagcaacaactacc 1001
Qy      2081  TGTGCTGCACTGCCCCAGGGGCAATC---GGGATGGGCCCCCAGAGTGAATTTCC 2137
Db      1002  agtgcctgcgtacacatgacacgtctcagaaagatgtgtgtgtgtgtgtgtgtccc 1061
Qy      2138  TCGATCTGACCTCCGCTTCAAGAGAGCTTGGCGAGGGCGAGTTGGGAGGTGACCT 2197
Db      1062  cagaaactcctactcactcaagaagcgtgagaaagacagatttgaggaggtcatct 1121
Qy      2198  GTGTGAGTGGAGACGCCCTCAAGATCGGTGACATGATTCCCTTATGTGCGTAA 2257
Db      1122  ctgtgaagtgagaggaatggaataatcaagaagatttgcctcagatgtcagtgcc 1181
Qy      2258  GGGACACCTTTGCTGTGCTGCTCAAGATCTTACGGCGAGATGCCACCAAGATGCGAG 2317
Db      1182  caacagcctgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1237
Qy      2318  CTCTCTCTTCTTCCAGAGATGATTTCTGAAAGAGGTGAAGATCATGTGAGGCTCAA 2377
Db      1238  -----ccagaaatgatttctcctaagagataagatcatgtctgcgtcaa 1283
Qy      2378  GGAOCCCAACATGATTTGGGCTGCTGGGCGTGTGTGTGACAGACACCCCTCTCATGAT 2437
Db      1284  ggaacaaacatacatcatatcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1343
Qy      2438  TACTGATACATGAGAGAGCGGACCTCAACAGATCTCTGAGTCCACACCGTGGAGGA 2497
Db      1344  caatgaataacatgtgagaaatgtgaaatcactcctcctccgcagagagccctaa 1403
Qy      2498  CAAGGACCGGAGGGGGCCCCGAGGACGGGACGCGAGGGGCGCCACCATCAGCTA 2557
Db      1404  ttcttcctccagcg-----atgtagcaactgtcagttta 1456
Qy      2558  CCCAATGCTCTGCAATGTGCGACGCCAGATCGCTCCGCAATGCGCATGTGCGCACAT 2617
Db      1437  caccacatctaaattatgtaccacaattgctcctgagatgaaagtaactctcctct 1496
Qy      2618  CAACCTTTGATCGGAGACCTGGGACGCGGAACTGCTTGTGGGAAATTCACCAT 2677
Db      1497  taatttgcacacgagatctgtgcacaaacgtctttagtggtaagaactacacaaat 1556
Qy      2678  CAAATGCGAGAGCTTGGCATGAGCCGGAACCTCTATGCTGGGAGCTATTAACCTGTGA 2737
Db      1557  caagatagcgtacttggaaatgagcagaaactgtacagatgagtgatcattacagca 1616
Qy      2738  GGGCGGGGAGTGTCTGCCATCCGCTGTGAGTGGCTGGAGATGCACTCTCATGGGAAAT 2797
Db      1617  gggccggagcagtgctccctacgtcgtgagatgtctgtgagaaatgtctgtcgtg 1676
Qy      2798  CAGGACTGCGAGTGTGAGTGGGCTTGTGATGACCTGTGGAGAGTGTGAGTGTGCTG 2857
Db      1677  caactaagcagagatgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1736
Qy      2858  TAGGGCCAGACCTTTGGGAGCTCAACGAGAGAGAGATGAGAGAGCGGGGAGTT 2917
Db      1737  tcaagaaagcgcctatccacagctgtcagatgaaacagtgatgagataactgtgag 1796
Qy      2918  CTTCGGGAGACAGGCGGAGAGTGTACCTGTCCGGCCCGCTGCTCCCGCAGAGGCT 2977
Db      1797  ctcccgagaccaaaggagagacttaactcctcaacagcattgtctcgtactgt 1856
Qy      2978  ATATGAGCTGATGCTGTGAGTGGAGCGGAGTGTGAGCAGAGAGACACCTTTTCCA 3037

```

```

Db      1857  gtaaaagctgatagtcagcgtcgtgagagagatacagaagaccgtccctcattccaaga 1916
Qy      3038  GCTGCATGCTGTTCC 3052
Db      1917  aatccacctctgtc 1931

RESULT 15
AAC00624
ID   AAC00624 standard; cDNA; 408 BP.
XX
AC   AAC00624;
XX
DT   06-OCT-2000 (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 622.
DE
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999; 99US-0122487.
XX
PA   (GIST ) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI: 2000-500381/45.
XX
PT   P-PSDB: AAG00618.
XX
PS   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PS   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PS   diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS   Claim 1; SEQ ID 622; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. An ORF has been identified within the
CC   sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors.
XX
SQ   Sequence 408 BP: 71 A; 113 C; 126 G; 85 T; 13 other;

Query Match      8.2%; Score 324.4; DB 21; Length 408;
Best Local Similarity 95.3%; Pred. No. 9.1e-62;
Matches 344; Conservative 8; Mismatches 7; Indels 2; Gaps 2;
Qy      278  GAGATGCTGCCCCACCCCTTAGGCCCCGAGGATCAAGAGCTATGGACCAAGGCCCT 337
Db      50  gagatgtgcgtcccccaccccttagggccgagagatcagagatgagacagagccct 109
Qy      338  GTCACTCTTACTGTGCTGCTGTGTGTCAGCAAGTGAAGATGACATGAAGGACATTT 397
Db      110  gtacatttactgtctgctgtctgtgtgtgcaagtgtgagatgctcgaatgagagacatt 169

```

